Journal of Hydrology 560 (2018) 326–341

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Impacts of correcting the inter-variable correlation of climate model outputs on hydrological modeling

Jie Chen [a,*], Chao Li [b], François P. Brissette [c], Hua Chen [a], Mingna Wang [d], Gilles R.C. Essou [c]

[a] State Key Laboratory of Water Resources & Hydropower Engineering Science, Wuhan University, 299 Bayi Road, Wuchang Distinct, Wuhan, Hubei 430072, China
[b] Department of Global Ecology, Carnegie Institution for Science at Stanford University, Stanford, CA, USA
[c] Department of Construction Engineering, École de Technologie Supérieure, Université du Québec, 1100 rue Notre-Dame Ouest, Montreal, QC H3C 1K3, Canada
[d] Department of Water Resources, China Institute of Water Resources and Hydropower Research, Beijing, China

ARTICLE INFO

Article history:
Received 3 February 2016
Received in revised form 21 September 2017
Accepted 15 March 2018
Available online 17 March 2018
This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Ralf Merz, Associate Editor

Keywords:
Climate change impacts
Bias correction
Inter-variable correlation
Climate model
Hydrology

ABSTRACT

Bias correction is usually implemented prior to using climate model outputs for impact studies. However, bias correction methods that are commonly used treat climate variables independently and often ignore inter-variable dependencies. The effects of ignoring such dependencies on impact studies need to be investigated. This study aims to assess the impacts of correcting the inter-variable correlation of climate model outputs on hydrological modeling. To this end, a joint bias correction (JBC) method which corrects the joint distribution of two variables as a whole is compared with an independent bias correction (IBC) method; this is considered in terms of correcting simulations of precipitation and temperature from 26 climate models for hydrological modeling over 12 watersheds located in various climate regimes. The results show that the simulated precipitation and temperature are considerably biased not only in the individual distributions, but also in their correlations, which in turn result in biased hydrological simulations. In addition to reducing the biases of the individual characteristics of precipitation and temperature, the JBC method can also reduce the bias in precipitation-temperature (P-T) correlations. In terms of hydrological modeling, the JBC method performs significantly better than the IBC method for 11 out of the 12 watersheds over the calibration period. For the validation period, the advantages of the JBC method are greatly reduced as the performance becomes dependent on the watershed, GCM and hydrological metric considered. For arid/tropical and snowfall-rainfall-mixed watersheds, JBC performs better than IBC. For snowfall- or rainfall-dominated watersheds, however, the two methods behave similarly, with IBC performing somewhat better than JBC. Overall, the results emphasize the advantages of correcting the P-T correlation when using climate model-simulated precipitation and temperature to assess the impact of climate change on watershed hydrology. However, a thorough validation and a comparison with other methods are recommended before using the JBC method, since it may perform worse than the IBC method for some cases due to bias nonstationarity of climate model outputs.

1. Introduction

The outputs of general circulation models (GCMs) have commonly been used to assess the impacts of climate change (e.g., Wilby and Harris, 2006; Sharma et al., 2007; Minville et al., 2008; Chen et al., 2011a,b). However, outputs from state-of-the-art climate models are usually too biased to be used as direct inputs to impact models. Climate model output bias is reflected in multiple aspects, such as the mean, standard deviation and inter-variable correlation. To address this issue in part, several bias correction methods have been developed to correct the critical characteristics of climate model outputs (e.g., mean, standard deviation). In particular, the distribution mapping methods have been found to be better than the mean-based methods for both climate simulation and hydrological modeling (Teutschbein and Seibert, 2012, 2013; Chen et al., 2013a). Bias correction methods were initially developed to correct biases of RCMs, but they are now widely used for GCMs (e.g. Li et al., 2014; Cannon, 2016; Chen et al., 2016, 2017) without a downscaling step. In such cases the bias correction accounts for the GCM biases and the difference between the GCM and local scales.

However, the commonly used bias correction methods correct climate variables independently. In other words, the interactions among climate variables are not specifically corrected. As a result, independent bias correction can produce non-physical corrections

* Corresponding author.
E-mail address: jiechen@whu.edu.cn (J. Chen).

https://doi.org/10.1016/j.jhydrol.2018.03.040
0022-1694/

and may fail to capture critical multivariate relationships. The misrepresented inter-variable correlations may lead to biased results in climate change impacts; in a likely scenario, the lack of inherent correlations between precipitation and temperature could result in biased hydrological simulations (Buishand and Brandsma, 2001; Maurer et al., 2010; Mueller and Seneviratne, 2014; Immerzeel et al., 2014). This is not surprising because the precipitation-temperature (P-T) correlations may affect evapotranspiration, snowmelt, and possibly runoff generation (Li et al., 2014).

Even though biases in the inter-variable correlations of climate model outputs have been pointed out by several studies, only a few of them have corrected correlation bias prior to use in climate change impact studies. For example, Thrasher et al. (2012) applied a quantile-mapping method for correcting daily maximum and minimum temperatures. This method first corrects both the maximum (or minimum) temperature and the diurnal temperature range, and then calculates the minimum (or maximum) temperature by subtracting (or adding) the corrected diurnal range. However, this approach is not suitable for climate variables, such as precipitation and temperature, which do not have deterministic relationships. In addition, Piani and Haerter (2012) proposed a two-dimensional bias correction method for correcting P-T correlation through the following three steps: 1) Correct the temperature; 2) Group precipitation and temperature pairs into a number of bins, and 3) Correct the precipitation within each temperature bin. Since the P-T correlation was not specifically considered, the benefits of this method could not be determined by assessing the P-T correlation coefficients. Li et al. (2014) also proposed a method for specifically correcting the joint distribution of GCM-simulated precipitation and temperature. This method can be seen as a bivariate extension of the commonly used univariate distribution mapping method, and has been found to be satisfactory in reducing the bias of GCM-simulated P-T correlation fields (Cannon et al., 2015; Gennaretti et al., 2015). More recently, Vrac and Friederichs (2015) and Cannon (2016) also proposed multivariate methods to correct climate model outputs. For example, Vrac and Friederichs (2015) proposed a copula-based bias correction method combining univariate bias correction with a shuffling method reproducing the empirical multivariate copula. Cannon (2016) proposed a multivariate analog of the quantile mapping method for correcting the correlation dependence structure of climate variables.

One of the ultimate goals of using bias-corrected climate simulations is to assess the climate change impacts, for example, on hydrology. While it may appear reasonable to conclude that limited benefits will be translated to the hydrological variables, small changes in P and T can sometimes results in much larger changes once the non-linear rainfall runoff filter has been applied. Since most hydrological climate change impact studies have not taken into consideration the inter-variable dependence, the potential impacts of this omission should be studied. There are currently no study that investigates the advantages of using multivariate bias correction for hydrological climate change impact studies. If the benefits of multivariate method are translated to hydrological variables, are they consistent across watersheds, climate zones or hydrological metrics? Accordingly, this paper aims at answering these questions by evaluating the effect of simultaneously correcting multiple variables, including precipitation and temperature, on hydrological modeling. To this end, we first correct the simulated precipitation and temperature simultaneously, using the joint bias correction (JBC) method of Li et al. (2014). Since the JBC method can be considered as a bivariate extension of the commonly used univariate distribution mapping method, it is feasible to compare the independent distribution mapping method in terms of hydrological modeling. We then compare the impacts of JBC on hydrological modeling over 12 watersheds covering various climate conditions with those of the commonly used distribution mapping method, which corrects precipitation and temperature independently. Since both methods only differ in correcting inter-variable correlations, any difference in hydrological simulations can be directly attributed to inter-variable correlation.

2. Study area and data

2.1. Study area

This study was conducted over 12 watersheds located in four countries across the globe (Table 1). Ten of the watersheds located in the United States (seven) and Canada (three) were also used in a previous study by Chen et al. (2013a) for assessing the performance of various bias correction methods. These ten watersheds were selected from different climate zones and topographic conditions. In terms of climate conditions, the 10 watersheds can be classified into five snowfall-dominated (Manic-5 (# 1), Carrot (# 2), Nation (# 3), Yampa (# 9) and Yellowstone (# 10)), three rainfall-dominated (Chickasawhay (# 4), Grand (# 5), and Umpqua (# 7)) and two rainfall-snowfall-mixed (Sacandaga (# 6) and Wolf (# 8)) watersheds. Although the three watersheds in Canada (Manic-5, Carrot, and Nation) are all snow-dominated, they have very different annual precipitation and temperature rates. Manic-5 is the coldest watershed, with abundant precipitation, while Carrot and Nation are located in arid regions. Three of the United States watersheds (Umpqua, Yampa and Yellowstone) are located in westerly mountainous areas, and the other four (Chickasawhay, Grand, Sacandaga and Wolf), in the centrally and easterly flatter terrain. The hydrological regime of the Umpqua can be divided into contrasting wet and dry seasons. Two watersheds (Xiangjiang (# 11) and Oueme (# 12)) with different climates (subtropical

Table 1
Basic information on the 12 selected watersheds (PCP = mean annual precipitation, and TMP = daily mean temperature).

Country	Watershed name	**Watershed ID**	Area (km^2)	PCP (mm)	TMP (°C)	Köppen-Geiger Climate classification		
						Main climate	PCP	TMP
Canada	Manic-5	**1**	24,610	950	−2.7	Snow	Fully humid	Cool summer
	Carrot	**2**	12,600	425	0.3	Snow	Fully humid	Cool summer
	Nation	**3**	6790	621	1.0	Snow	Fully humid	Cool summer
United States	Chickasawhay	**4**	6967	1477	17.6	Warm temperate	Fully humid	Hot summer
	Grand	**5**	5825	919	10.9	Snow	Fully humid	Warm summer
	Sacandaga	**6**	1271	1260	5.1	Snow	Fully humid	Warm summer
	Umpqua	**7**	9535	1356	8.7	Warm temperate	Summer dry	Warm summer
	Wolf	**8**	5851	799	5.6	Snow	Fully humid	Warm summer
	Yampa	**9**	8828	655	3.3	Snow	Fully humid	Warm summer
	Yellowstone	**10**	6791	929	−0.7	Snow	Fully humid	Cool summer
China	Xiangjiang	**11**	52,150	1571	18.8	Warm temperate	Fully humid	Hot summer
Benin	Oueme	**12**	50,000	1149	27.2	Warm temperate	Fully humid	

monsoon and tropical) were also selected from China and Benin, respectively. The Xiangjiang watershed, located in the central south region of China, is a large sub-basin of the Yangtze River watershed. It has a subtropical monsoon climate, and a mean annual precipitation of about 1570 mm. Most of the precipitation occurs between April and June, resulting in high flows over this period, while little occurs in the winter, which is defined as a low-flow period. The Oueme watershed located in the Republic of Benin, West Africa is characterized by two different climates: a subequatorial climate to the south and a Sudanian climate to the north. The temperature is high all the year round, and so there is no snow occurring in the winter. Heavy rainfall mainly occurs between April and October (accounts for 92%), with minimal amounts during the other months.

2.2. Data

The observed meteorological data (daily precipitation and temperature) for the three Canadian watersheds studied were drawn from the Hutchinson et al. (2009) gridded dataset, which was obtained by interpolating station data to a 10 km grid using a thin plate smoothing spline surface fitting method. For the seven watersheds in the United States, the 5 km Santa Clara dataset was used (Maurer et al., 2002; Livneh et al., 2013). To run a lumped hydrological model, as will be introduced section 3.b, all grid points within the watersheds were averaged to one for both the precipitation and temperature time series. The gauge-measured precipitation and temperature time series were used for the Xiangjiang and Oueme watersheds. The mean precipitation and temperature of the Xiangjiang watershed were obtained by averaging the precipitation and temperature over 100 rain gauges and 8 temperature gauges within or surrounding the watershed, respectively, based on weights calculated using the Thiessen polygon method (Rhynsburger, 1973). The same averaging method was also used for the Oueme watershed, based on 18 rain gauges and 3 temperature gauges.

Discharge data for Manic-5 (# 1) were provided by Hydro-Québec, drawn from mass balance calculations at the dam. For the other two Canadian watersheds (Carrot (# 2) and Nation (# 3)), they were obtained from Environment Canada. The discharge data for the United States watersheds were obtained from the Model Parameter Estimation Experiment (MOPEX) database (Duan et al., 2006). Discharge data for Xiangjiang and Oueme were respectively obtained from the Department of Water Resource of Hunan Province, China and Direction Générale de l'Eau (DG-Eau), Benin.

The climate model simulations were taken from the Coupled Model Intercomparison Project Phase 5 (CMIP5) database. Twenty-six GCMs (ID 1–26 in Table 2) from 16 modeling centres were selected to adequately represent the climate model uncertainty (Table 2). Depending on their surface areas, some watersheds may include only one or even no GCM grid box. To obtain the mean watershed precipitation and temperature, data from several GCM grids (more than four, depending on the size of the watersheds) within or surrounding the watersheds were averaged using the Thiessen polygon method. The dataset covers the period of 1975–2004 for the Xiangjiang watershed, and of 1961–2000 for the other 11 watersheds. The first half of each period was used to calibrate the bias correction methods, and the second half for validation.

3. Methodology

3.1. Bias correction methods

The JBC method of Li et al. (2014) was used to jointly correct precipitation and temperature simulated by 26 GCMs. Since the

Table 2
Basic information on the 26 selected general circulation models (GCMs) from CMIP5 database.

GCM ID	GCM name	Modeling centre	Institution	Horizontal resolution (lon. × lat.)
1	ACCESS1.0	CSIRO-BOM	CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia), and BOM (Bureau of Meteorology, Australia)	1.875 × 1.25
2	ACCESS1.3			1.875 × 1.25
3	BCC-CSM1.1(m)	BCC	Beijing Climate Centre, China Meteorological Administration	1.125 × 1.125
4	BCC-CSM1.1			2.8 × 2.8
5	BNU-ESM	GCESS	College of Global Change and Earth System Science, Beijing Normal University	2.8 × 2.8
6	CanCM4	CCCma	Canadian Centre for Climate Modelling and Analysis	2.8 × 2.8
7	CanESM2			2.8 × 2.8
8	CMCC-CESM	CMCC	Centro Euro-Mediterraneo per I Cambiamenti Climatici	3.75 × 3.7
9	CMCC-CMS			1.875 × 1.875
10	CNRM-CM5	CNRM-CERFACS	Centre National de Recherches Meteorologiques/Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	1.4 × 1.4
11	CSIRO-Mk3.6.0	CSIRO-QCCCE	Commonwealth Scientific and Industrial Research Organisation in collaboration with the Queensland Climate Change Centre of Excellence	1.8 × 1.8
12	FGOALS-g2	LASG-CESS	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences; and CESS, Tsinghua University	1.875 × 1.25
13	GFDL-CM3	NOAA-GFDL	Geophysical Fluid Dynamics Laboratory	2.5 × 2.0
14	GFDL-ESM2G			2.5 × 2.0
15	GFDL-ESM2M			2.5 × 2.0
16	GISS-E2-H	NASA-GISS	NASA Goddard Institute for Space Studies	2.5 × 2.0
17	GISS-E2-R			2.5 × 2.0
18	INMCM4	INM	Institute for Numerical Mathematics	2.0 × 1.5
19	IPSL-CM5A-LR	IPSL	Institut Pierre-Simon Laplace	3.75 × 1.8
20	IPSL-CM5A-MR			2.5 × 1.25
21	IPSL-CM5B-LR			3.75 × 1.8
22	MIROC5	MIROC	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	1.4 × 1.4
23	MIROC-ESM-CHEM	MIROC	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8 × 2.8
24	MIROC-ESM			2.8 × 2.8
25	MRI-CGCM3	MRI	Meteorological Research Institute	1.1 × 1.1
26	MRI-ESM1			1.125 × 1.125

original JBC method was developed for bivariate correction, only daily precipitation and mean temperature were considered. The daily mean temperature was obtained by averaging the daily maximum and minimum temperatures for both observed and GCM-simulated data. A previous study by Chen and Brissette (2015) showed that daily mean temperature estimated using this method is as good as that estimated by averaging hourly temperature (24 values per day). To investigate the benefits of the joint correction of precipitation and temperature on hydrological modeling, the results obtained from JBC were compared with those obtained from an independent bias correction (IBC) method. Prior to using JBC and IBC methods, the wet-day frequency of GCM-simulated precipitation was corrected using the local intensity scaling (LOCI) method of Schmidli et al. (2006). With the LOCI method, a wet-day threshold was first determined using the GCM daily precipitation time series for each month such that the threshold exceedance matched the wet-day frequency of the observed time series for the calibration period. This threshold was then used to correct the wet-day frequency of GCM-simulated precipitation for the validation period. This threshold value varies from one GCM to the other and ranged between 0 and 6.6 mm. After the wet-day frequency correction, precipitation amounts were respectively corrected using the IBC and JBC methods. The IBC method used in this study is a distribution mapping method with a gamma distribution modeling precipitation and Gaussian distribution for temperature. This method has been used in several other studies (e.g., Chen et al., 2013a,b, Ines and Hansen, 2006; Piani et al., 2010) to produce climate simulation for hydrological modeling and climate change impact studies. The JBC method is a bivariate extension of the IBC method, and it is only briefly described here. More details can be found in Li et al. (2014). When using the JBC method, a joint cumulative distribution function (CDF) must first be built for precipitation and temperature. The gamma and Gaussian distributions are used to model the marginal distributions of precipitation and temperature, respectively. With the joint CDF, the marginal and conditional distributions are then derived, which are then used to sequentially correct the precipitation and temperature. JBC can be performed by first correcting either the precipitation or the temperature. In this study, we start by correcting the precipitation, and then go on to the temperature. Both IBC and JBC were conducted at the monthly basis.

3.2. Hydrology modeling

The hydrological modeling was carried out using a lumped, conceptual, rainfall-runoff model, HMETS, developed by École de technologie supérieure. The model structure is detailed in Martel et al., (2017). The HMETS source code, interface and user guide are freely available on the MathWorks file exchange site. HMETS has been used in several flow forecasting and climate change impact studies (e.g., Chen et al., 2011b; Essou and Brissette, 2013; Arsenault et al., 2015, 2017; Arsenault and Brissette, 2016). The model has up to 21 free parameters: one for evapotranspiration, ten for snow accumulation and snowmelt, four for vertical water balance, and six for horizontal water movement. Information about each parameter is presented in Table 3.

The calculation of potential evapotranspiration (PET) follows the work of Oudin et al. (2005). A constant parameter is used during optimization to adjust PET to a specific watershed. Actual evapotranspiration depends on water present in the vadose zone reservoir of the hydrological model. The snowmelt model is based on the work of Vehvilainen (1992), which is a degree-day model that allows for melting and refreezing process within the snowpack. The model works in three steps including the overnight refreezing process, snowmelt and snowpack water retention capacity. The vertical water balance takes into account all the exchanges made between the surface, vadose and saturated zones. Depending on the water level in the vadose zone reservoir, available surface water is partitioned between runoff and infiltration once the actual evapotranspiration has been taken out. The saturated zone is represented by a linear reservoir releasing groundwater flow. Outlet streamflow is calculated based on four horizontal transfer components. For surface and delayed runoff, two unit hydrographs based on the two-parameter gamma distribution density function are used to transfer water at the outlet. Total outlet streamflow is ultimately computed by summing up all four horizontal flow components. For purely rainfall-driven watersheds (e.g. Oueme (# 12)), the HMETS snow module was not used and the model reverted to its 11-parameter version. Model calibration was done automatically using the Covariance Matrix Adaptation Evolution Strategy (CMAES) (Hansen and Ostermeier, 1996, 2001), which was shown to be most efficient for this type of optimization class (Arsenault et al., 2013).

HMETS requires the daily mean temperature and precipitation as inputs. If the maximum and minimum temperatures are input to the model, they are automatically averaged to the mean temperature. A natural flow or discharge time series is also needed for model calibration and validation. HMETS structure is fairly typical of this class of hydrological models.

The use of a lumped conceptual model was based upon several considerations. Even though distributed hydrological models do have advantages in studying the spatial variability of hydrological response to climate and land use changes, when the goal is streamflow simulation/prediction at watershed outlets, lumped models have been shown to be just as good as their more complex distributed counterparts in several inter-comparison studies (e.g. Reed et al., 2004; Smith et al., 2012). Streamflow generation at a watershed outlet is the result of various spatially-integrated

Table 3
List of HMETS 21 parameters.

Parameter	Definition
Real evapotranspiration (1 parameter)	
ET_{eff}	Fraction of the potential evapotranspiration
Snowmelt model parameters (10 parameters)	
ddf_{min}	Minimum degree-day-factor in mm/°C/day
ddf_{plus}	Maximum degree-day-factor in mm/°C/day ($ddf_{min} + ddf_{plus} = ddf_{max}$)
T_{bm}	Base melting temperature in °C
K_{cum}	Empirical parameter for the calculation of the degree-day-factor in mm^{-1}
fc_{min}	Minimum fraction for the snowpack water retention capacity
fc_{plus}	Maximum fraction for the snowpack water retention capacity ($fc_{min} + fc_{plus} = fc_{max}$)
C_{cum}	Parameter for the calculation of water retention capacity in mm^{-1}
T_{bf}	Base refreezing temperature in °C
K_f	Degree-day factor for refreezing in mm/°C/day
Fe	Empirical exponent for the freezing equation
Subsurface (6 parameters)	
c_r	Fraction of the water for surface and delayed runoff
c_{vp}	Fraction of the water for groundwater recharge
c_v	Fraction of the water for hypodermic flow
c_p	Fraction of the water for groundwater flow
LV_{max}	Maximum level of the vadose zone in mm
LP_{max}	Maximum level of the phreatic zone in mm
Unit hydrograph parameters (4 parameters)	
α_1	Shape parameter α for the gamma distribution used on the surface unit hydrograph
β_1	Rate parameter β for the gamma distribution used on the surface unit hydrograph
α_2	Shape parameter α for the gamma distribution used on the delayed unit hydrograph
β_2	Rate parameter β for the gamma distribution used on the delayed unit hydrograph

processes and this explains why lumped model excel at this task and can do very well on very large watersheds.

The observed precipitation, temperature and discharge were used to calibrate and validate HMETS for all 12 watersheds. To ensure a robust calibration, both the calibration and validation periods were longer than 10 years for all the watersheds. The optimal combination of parameters was chosen based on the Nash-Sutcliffe coefficient (NSE, Nash and Sutcliffe, 1970). The chosen set of parameters yielded NSE values ranging between 0.766 (Grand) and 0.924 (Xiangjiang) for calibration, and between 0.724 (Carrot) and 0.892 (Chickasawhay) for validation, across all 12 watersheds, indicating the good performance of HMETS for these watersheds (Table 4) and the suitability of using a lumped model for this study. The observed and simulated mean hydrographs presented in Fig. 1 also outline the good performance of HMETS for these watersheds.

Table 4
Nash-Sutcliffe coefficients (NSE) of the calibration and validation of hydrological model HMETS and the period used to calculate the high and low flows.

Watershed ID	Calibration period	NSE calibration	Validation period	NSE validation	High flow period	Low flow period
1	1966–1980	0.890	1981–1995	0.825	Apr-Jun	Jul-Nov
2	1967–1983	0.833	1984–2000	0.724	Apr-Jun	Jul-Nov
3	1981–1990	0.898	1991–2000	0.864	Apr-Jul	Aug-Nov
4	1961–1980	0.917	1981–2000	0.892	Dec-May	Jun-Nov
5	1961–1980	0.766	1981–2000	0.790	Jan-Jul	Aug-Dec
6	1961–1980	0.823	1981–2000	0.743	Mar-May	Jun-Feb
7	1961–1980	0.877	1981–2000	0.836	Nov-Mar	Apr-Oct
8	1961–1980	0.874	1981–2000	0.790	Mar-Jun	Jul-Nov
9	1961–1980	0.795	1981–2000	0.756	Apr-Jul	Aug-Nov
10	1961–1980	0.870	1981–2000	0.834	Apr-Jul	Aug-Nov
11	1975–1989	0.924	1990–2004	0.866	Mar-Jul	Aug-Feb
12	1961–1980	0.885	1981–2000	0.871	Jun-Nov	Dec-May

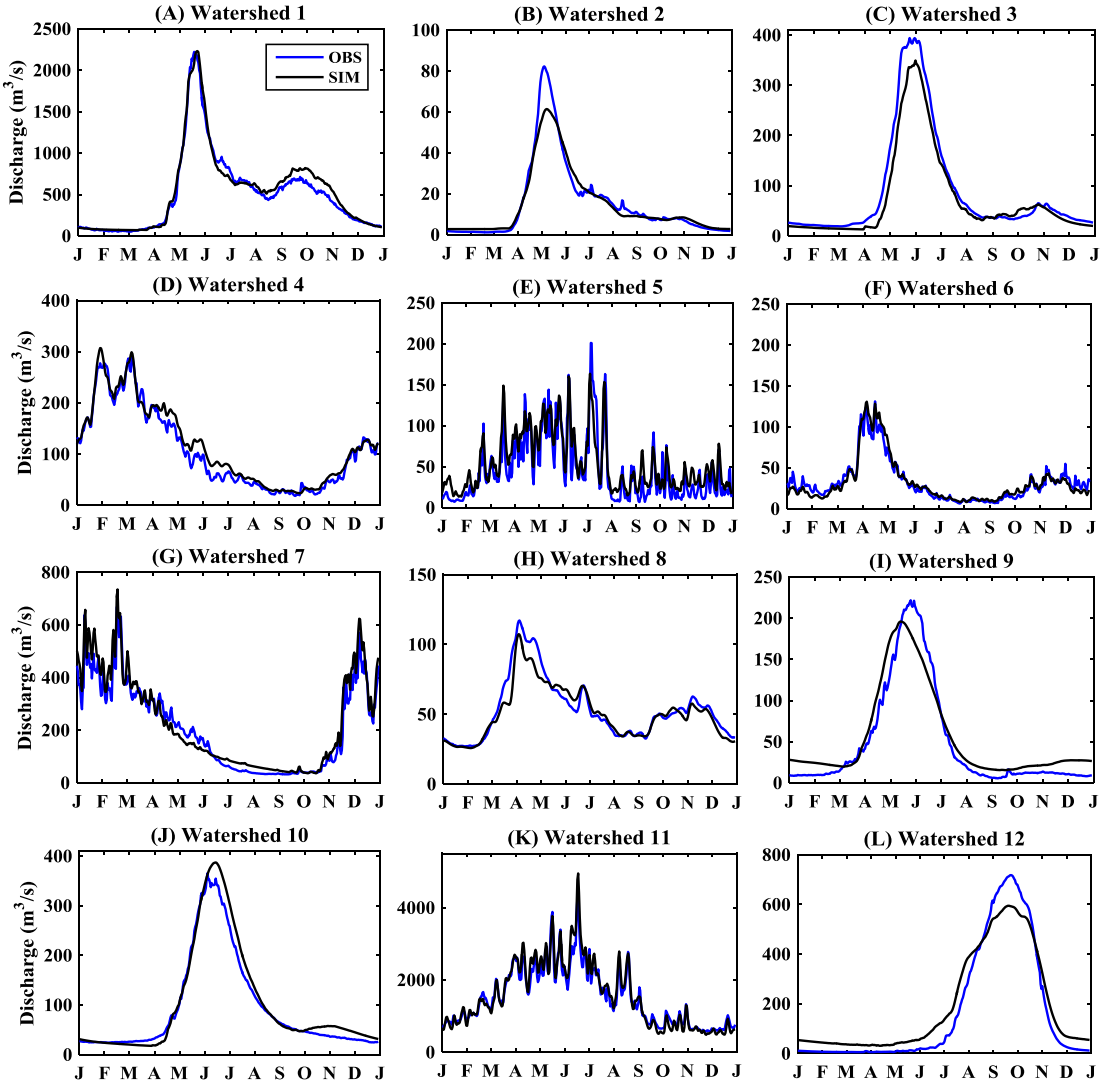

Fig. 1. Mean annual hydrographs of observed discharge (OBS) and observed precipitation and mean temperature simulated discharge (SIM) for the validation period.

3.3. Data analysis

This study first evaluated the performance of both bias correction methods (IBC and JBC) in correcting GCM-simulated precipitation and temperature, and then for hydrological modeling. The corrected results from both methods were each compared to the other, as well as with the raw GCM outputs, in order to show the benefits of bias correction. Since the performance of

Table 5
Metrics to evaluate the performance of independent and joint methods in correcting precipitation and temperature for representing watershed streamflow (STD = standard deviation).

ID	Mean	ID	STD	ID	Quantile	ID	Others
1	All days mean	14	All days STD	27	Quantile 0.01	38	Time to peak discharge
2	January mean	15	January STD	28	Quantile 0.10	39	Time to the beginning of flood
3	February mean	16	February STD	29	Quantile 0.20	40	Time to the end of flood
4	March mean	17	March STD	30	Quantile 0.30	41	10-year return period high flow
5	April mean	18	April STD	31	Quantile 0.40	42	20-year return period high flow
6	May mean	19	May STD	32	Quantile 0.50	43	50-year return period high flow
7	June mean	20	June STD	33	Quantile 0.60	44	10-year return period low flow
8	July mean	21	July STD	34	Quantile 0.70	45	20-year return period low flow
9	August mean	22	August STD	35	Quantile 0.80	46	50-year return period low flow
10	September mean	23	September STD	36	Quantile 0.90		
11	October mean	24	October STD	37	Quantile 0.99		
12	November mean	25	November STD				
13	December mean	26	December STD				

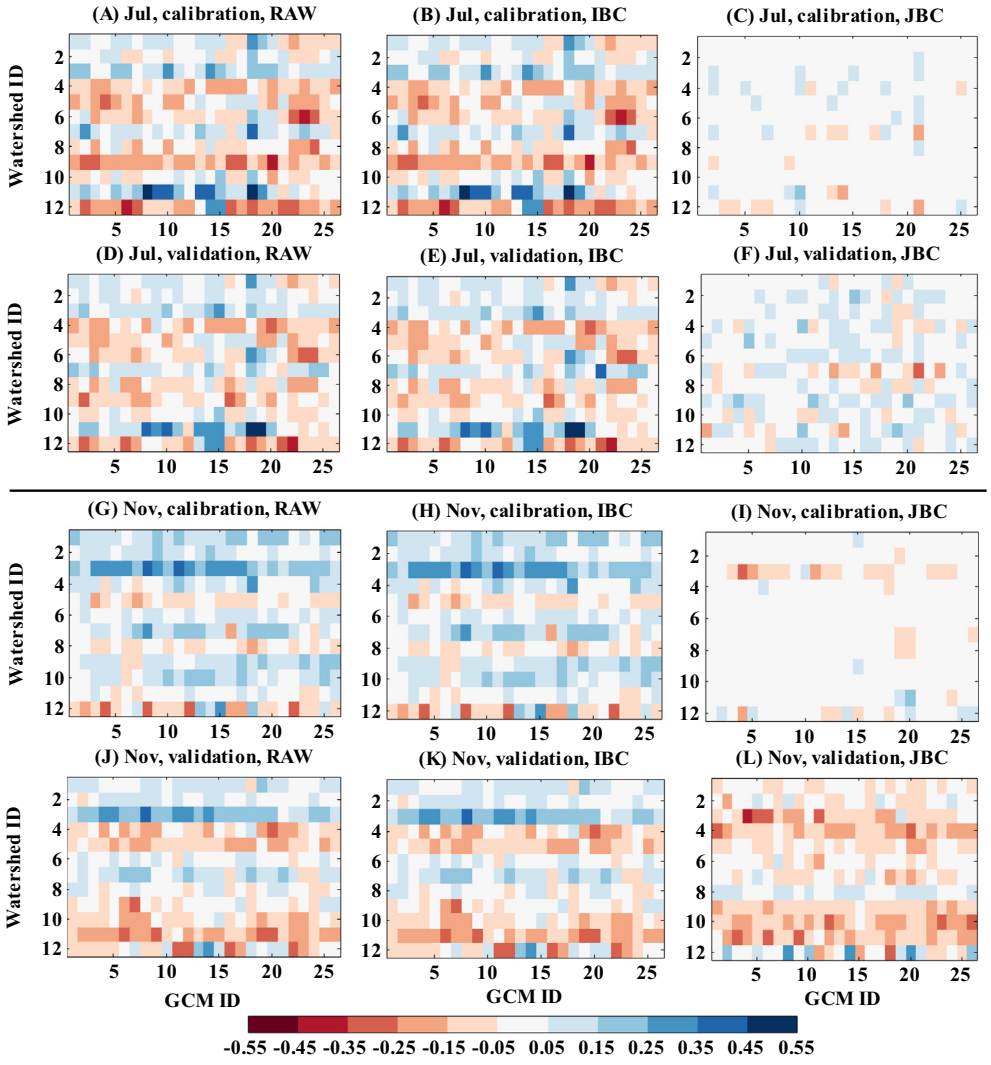

Fig. 2. Absolute error of the Kendall's tau correlation coefficient between precipitation and temperature of raw GCM-simulated (RAW) and independent (IBC) and joint (JBC) bias correction methods-corrected data for all 26 GCMs and 12 watersheds over both calibration and validation periods. Fig. 4(A)–(F) are results for July (Jul) and 4(G) to (L) are results for November (Nov). The watershed ID on the y-axis corresponds to the 12 watersheds, as presented in Table 1, and the GCM ID on the x-axis corresponds to the 26 GCMs, as presented in Table 2.

distribution mapping methods in terms of correcting individual variables has been thoroughly evaluated in several studies (e.g., Piani et al., 2010; Teutschbein and Seibert, 2012; Chen et al., 2013a; Li et al., 2014), this study only compares the performance of these two methods in terms of correcting P-T correlation. The absolute error (AE) was calculated for the P-T correlation coefficient for raw and corrected data (the difference between raw/corrected GCM and observed correlation coefficient). Since precipitation and temperature may not necessarily be linearly correlated, the Kendall's tau coefficient (Kendall, 1970) was used in this study.

To investigate the impacts of the JBC method on hydrological modeling, the raw and bias-corrected simulations of precipitation and temperature were used to drive HMETS for hydrological simulations over both the calibration and validation periods. Discharges simulated using IBC- and JBC-corrected simulations were compared with those simulated using observations according to 46 evaluation metrics (Table 5). These evaluation metrics capture the mean (13 metrics), standard deviation (13 metrics) and distribution (11 metrics) characteristics of the observed hydrology. The other 9 of the 46 metrics represent the timing variables (e.g., time to peak discharge (annual maximum flood), and time to the beginning and to the end of annual maximum flood) and the frequency of high and low flows (10-year, 20-year and 50-year return periods of high flow (0.95 quantile) and low flow (0.05 quantile). Time to the beginning and to the end of flood are determined based on a cumulative annual hydrograph (Chen et al., 2011b). Four breakpoints in the cumulative annual hydrograph are fitted and connected by straight lines, minimizing the root-mean-square error with the cumulative annual hydrograph. The time to the first breakpoint is defined as the beginning of flood, and the time to the second is defined as the end of flood. To obtain the return periods, the annual high and low flows are modeled using the log Pearson type-III distribution (USGS, 1982). The time periods used to calculate the high and low flows are watershed-dependent, as presented in Table 4.

The performance of the IBC and JBC methods with respect to reproducing climate simulations for hydrological modeling was evaluated with the following three steps: (1) The RE was first calculated for all 46 metrics and two bias correction methods; (2) The absolute value of the relative error (ARE) was then calculated for all metrics and both bias correction methods. The relative error (RE) was calculated for hydrological metrics as the difference between raw/corrected and observed data simulated values divided by the observed data simulated values; and (3) The difference in ARE (D_{ARE}) over two bias correction methods was calculated at the last step using Eq. (1).

$$D_{ARE} = |(M_{IBC} - M_{obs})/M_{obs}| - |(M_{JBC} - M_{obs})/M_{obs}| \quad (1)$$

where M represents a hydrological metric; $(M_{IBC}-M_{OBS})/M_{OBS}$ and $(M_{JBC}-M_{OBS})/M_{OBS}$ are REs of each hydrological metric simulated by

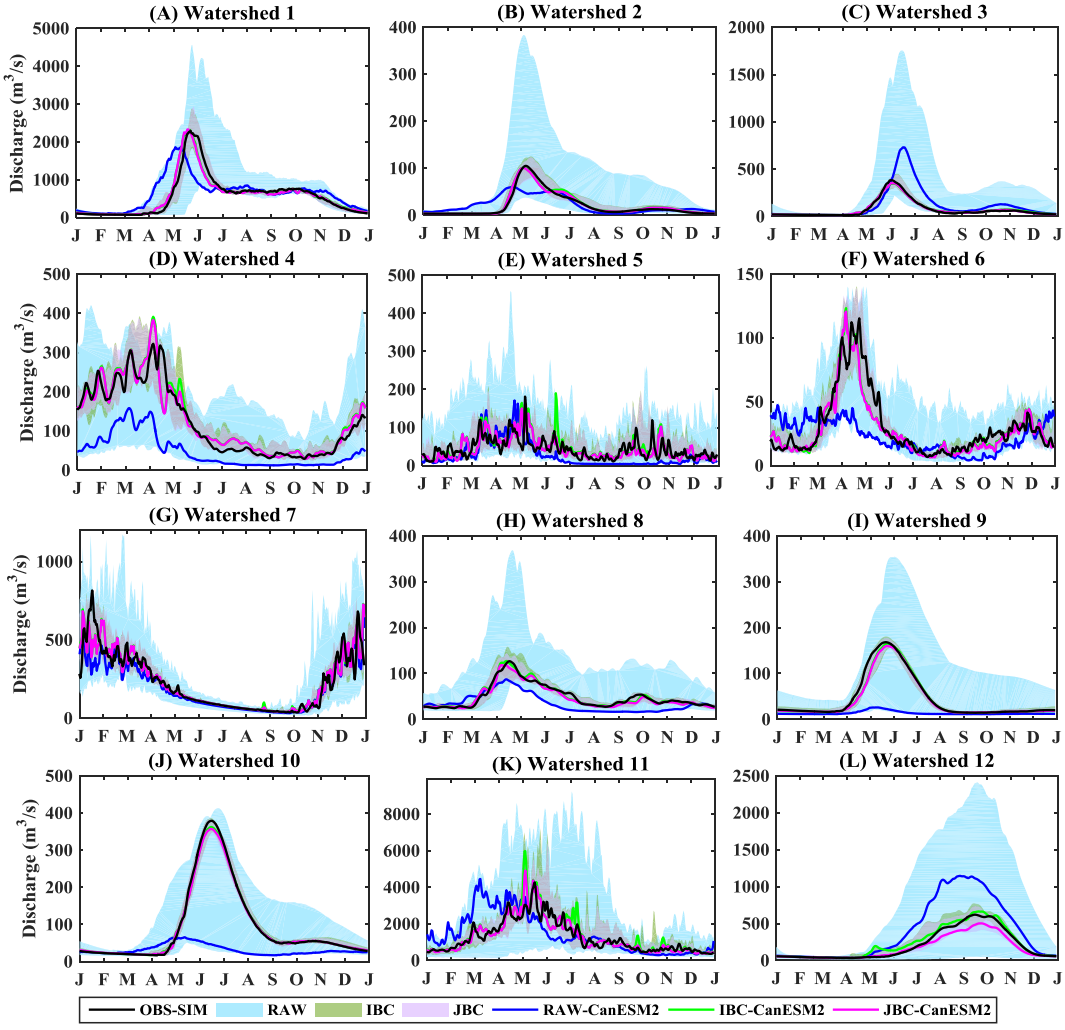

Fig. 3. Mean annual hydrographs simulated using observed (OBS-SIM), raw GCM-simulated (RAW), and independent (IBC) and joint (JBC) bias correction methods-corrected precipitation and temperature time series for all GCMs (results from the CanSEM2 simualtion are highlighted) and all 12 watersheds over the calibration period.

IBC- and JBC-corrected simulations, respectively. To avoid biases resulting from the hydrological modeling process, the observed streamflow time series is represented by modeled discharge using observed precipitation and temperature rather than the gauged observation. The D_{ARE} values were presented using portrait diagrams following the work of Gleckler et al. (2008).

The paired *t*-test was used to test the equality of the ARE of the IBC and JBC methods for all 46 metrics (46 pairs). A right-tailed paired *t*-test was first applied at a significance level of 5%. If the P-value of the paired *t*-test was <0.05, the JBC method is significantly better than the IBC method. In contrast, if the P-value is greater than 0.05, the performance of the JBC method is not statistically better than the IBC method. To test if the IBC method is better than the JBC method, a left-tailed paired *t*-test was further applied at the P = 0.05 level. If IBC is not statistically better than JBC, both methods are then considered as statistically identical. In addition, the mean absolute relative error (MARE) was calculated by averaging 26 (GCMs) × 46 (Metrics) AREs across all 26 GCMs and across all 46 hydrological metrics for both IBC- and JBC-corrected simulations over the calibration and validation periods. The relative difference of MARE (RD-MARE) between IBC- and JBC-corrected simulations was then calculated. A positive RD-MARE implies that the JBC method performs better than the IBC method and vice versa. Averaging AREs across GCMs reflects the overall performance of both bias correction methods in terms of hydrological metrics, while averaging AREs across all 46 metrics outlines the overall performance of both bias correction methods in terms of GCMs.

4. Results

4.1. Precipitation and temperature correlation

The AE of the P-T correlation coefficients of raw and IBC- and JBC-corrected data are calculated for all 26 GCMs and 12 watersheds over both the calibration and validation periods. Results of two typical months (July and November) are presented in Fig. 2 for illustration. The results show that all GCMs are considerably biased in preserving the observed P-T correlation. However, the sign and magnitude of biases are dependent on multiple factors: watershed location, GCM and month. For example, the P-T correlations for the three Canadian watersheds are overestimated by most GCMs in both July and November. However, they are underestimated for other watersheds in July. The IBC method correcting precipitation and temperature distributions independently retain the P-T dependencies of GCMs. The correlation coefficient is almost identical to that in raw GCM simulations. The JBC method performs reasonably well in reducing the bias of GCM P-T correlations for all months and watersheds over the calibration period. However, its

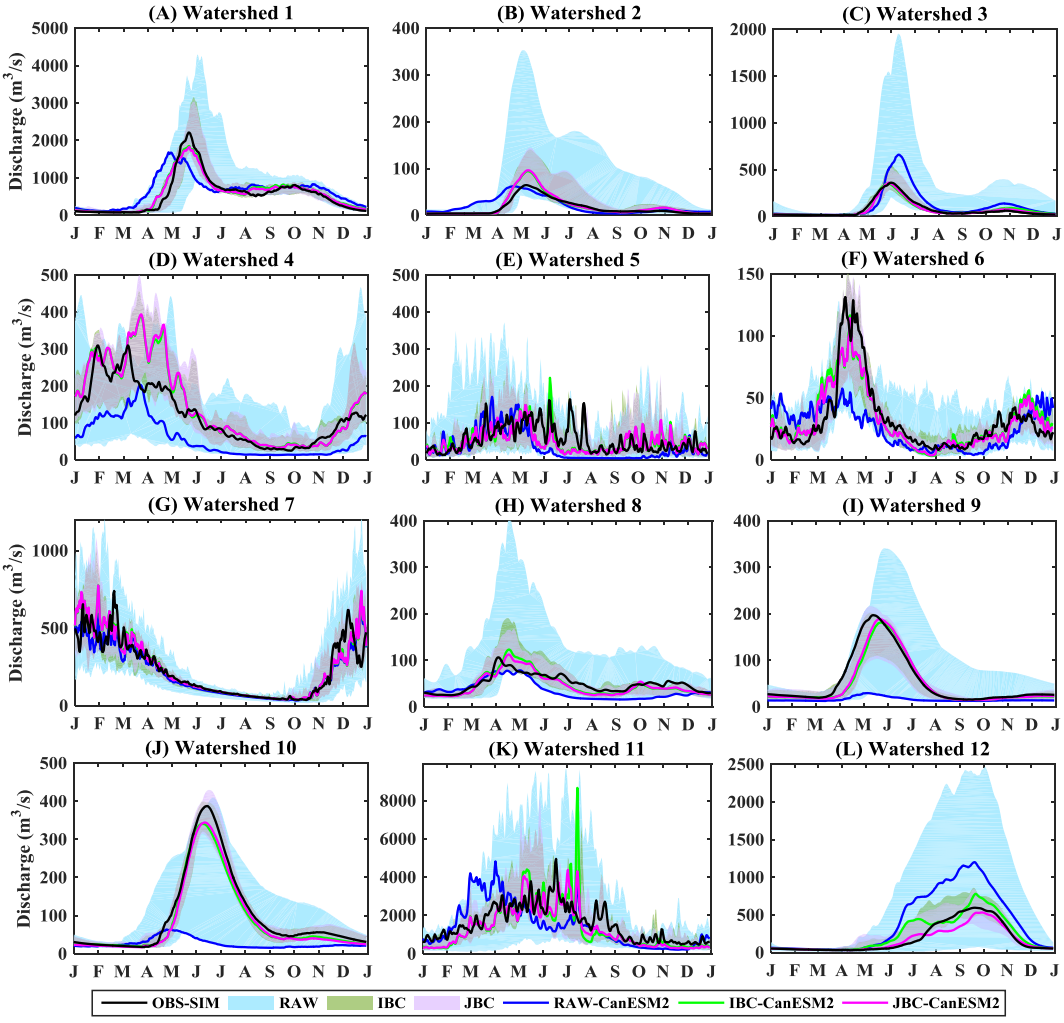

Fig. 4. Mean annual hydrographs simulated using observed (OBS-SIM), raw GCM-simulated (RAW), and independent (IBC) and joint (JBC) bias correction methods-corrected precipitation and temperature time series for all GCMs (results from the CanESM2 simulation are highlighted) 12 watersheds over the validation period.

performance in the validation period is monthly dependent. For example, the P-T correlation biases are considerably reduced for July, but not for November. Results for other months range somewhere between July and November.

4.2. Hydrological modeling

The raw and bias-corrected daily precipitation and temperature time series were used as inputs of HMETS to simulate streamflows over all 12 watersheds. Figs. 3 and 4 respectively present the mean hydrographs of all the watersheds over the calibration and validation periods. All GCMs results without (RAW) and with bias correction (IBC and JBC) are shown as envelopes and hydrographs simulated by CanESM2 simulations are presented as a typical example. The mean hydrographs simulated by the observed precipitation and temperature are also presented for comparison. Figs. 3 and 4 show that the mean hydrographs are poorly represented by raw GCM simulations, as shown by the wide envelope. In particular, the annual cycle of streamflows simulated by CanESM2 simulations cannot be captured for the Yampa (# 9) and Yellowstone (#10) watersheds. Both the IBC and JBC methods are capable of reducing the biases of GCM simulations and better representing the mean hydrograph over both the calibration and validation periods. However, when all simulations are plotted together, the difference between the IBC and JBC methods is difficult to notice, even though a slight advantage of the JBC method can be observed in some cases (e.g. Sacandaga (# 6) and Oueme (# 12) watersheds). In other words, both bias correction methods perform similarly with respect to reproducing mean hydrological conditions, as represented by the mean annual hydrograph.

In order to investigate the impacts of correcting GCM P-T correlation on hydrological modeling in greater detail, the IBC and JBC methods were further compared with respect to reproducing 46 hydrological metrics. Figs. 5 and 6 present the portrait diagrams of the D_{ARE} of 46 metrics for the calibration and validation periods, respectively. The warm color (red) in the portrait diagrams indicates that the IBC method performs better than the JBC method, while the cool color (blue) indicates that the JBC method is better. Overall, the JBC method performs better than the IBC method for most watersheds and GCMs over the calibration period, as illustrated by the fact that there are more "blue" than "red" indications in Fig. 5. The paired *t*-test shows that the JBC method performs significantly better than the IBC method for 11 out of the 12 watersheds at the 5% significance level (Table 6), indicating the

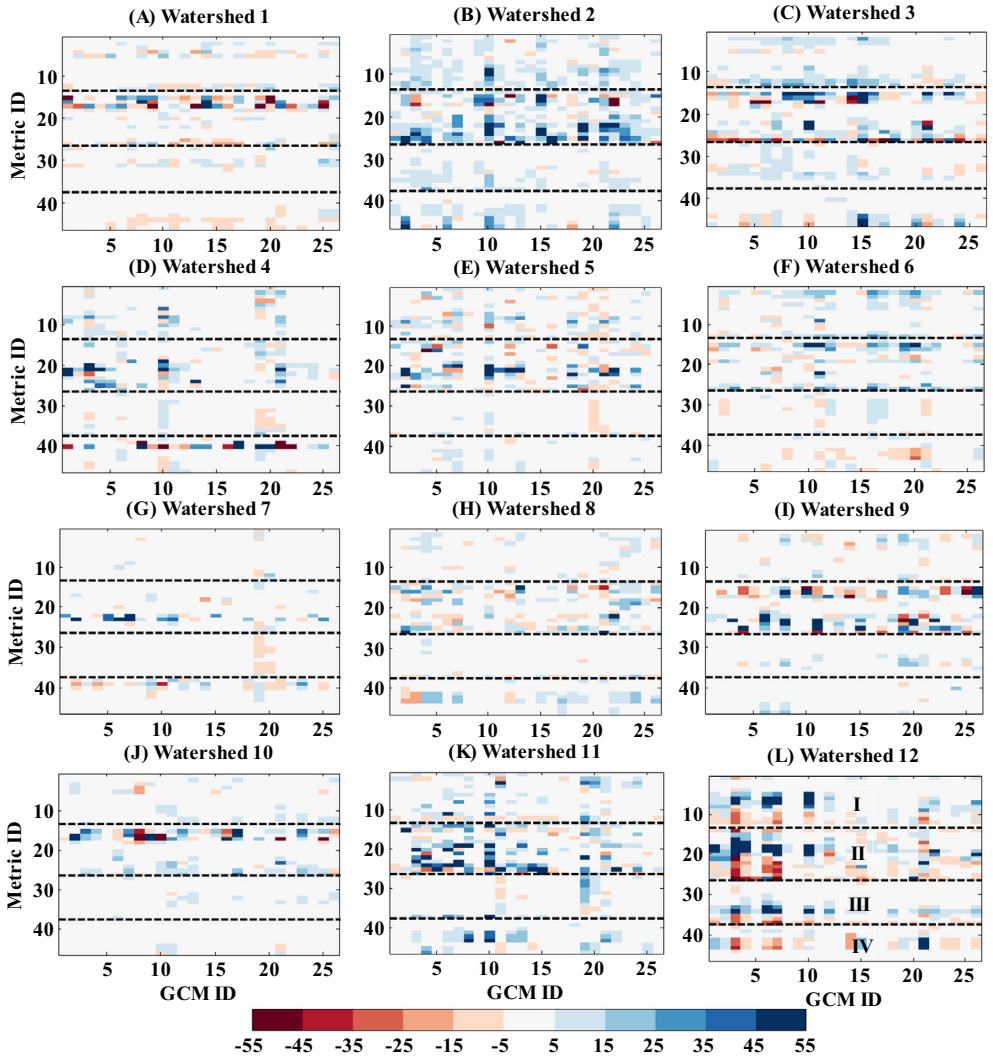

Fig. 5. Portrait diagram of the difference in absolute relative error (%) of the 46 metrics between streamflows simulated using independent (IBC) and joint (JBC) methods-corrected precipitation and temperature over the calibration period. I–IV in Fig. 7(L) represent mean, standard deviation, quantiles and other aspects (time variables and frequency of high and low flows) of daily streamflow. The metric ID on the y-axis corresponds to 46 hydrological metrics, as presented in Table 3, and the GCM ID on the x-axis corresponds to 26 GCMs, as presented in Table 2.

importance of correcting the P-T correlation of GCM simulations in hydrological impact studies (Table 6). However, for the Manic-5 watershed, both methods perform similarly. For the validation period, the portrait diagram mixes "blue" and "red" colors (Fig. 6). Over the validation period, it is hard to identify the advantage of one method over the other. The paired *t*-test shows that the JBC method performs significantly better than the IBC method for 6 out of the 12 watersheds at the 5% level. However, both methods show similar performance for 3 watersheds, and IBC performs significantly better than JBC for 3 watersheds.

The performance of both methods is also compared using MARE as a criterion. Figs. 7 and 8 show the RD-MARE between the IBC and JBC methods in terms of 26 GCMs over the calibration and validation periods. In line with results presented in Fig. 5, the RD-MARE shows that the JBC method consistently performs better than the IBC method for most GCMs over the calibration period. This applies to all watersheds with the exception of Manic-5 (# 1). In particular, the JBC method performs better for all GCMs over the Carrot watershed (# 2). For the validation period, the JBC method performs better than the IBC method for 6 out of 12

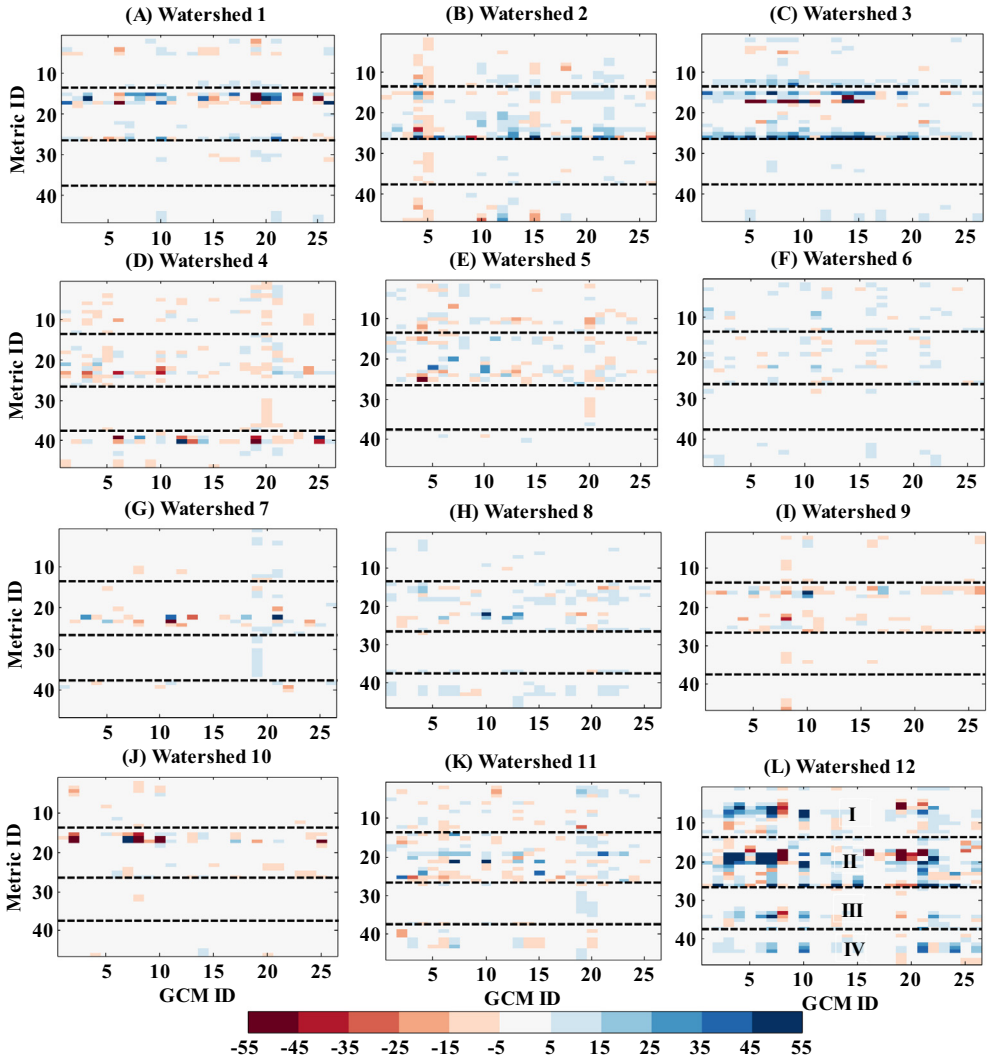

Fig. 6. Portrait diagram of the difference in absolute relative error (%) of 46 metrics between streamflows simulated using independent (IBC) and joint (JBC) methods-corrected precipitation and temperature over the validation period. I–IV in Fig. 8(L) represent mean, standard deviation, quantiles and other aspects (time variables and frequency of high and low flows) of daily streamflow. The metric ID on the y-axis corresponds to 46 hydrological metrics, as presented in Table 3, and the GCM ID on the x-axis corresponds to 26 GCMs, as presented in Table 2.

Table 6
P-values of the right/left tail paired *t*-test to test the equality of absolute relative errors of two bias correction methods in terms of 46 hydrological metrics for both calibration and validation periods (the green color denotes the significant difference at the P = 0.05 level).

Test	Watershed ID	1	2	3	4	5	6	7	8	9	10	11	12
Right tail test	Calibration	0.864	<0.001	<0.001	0.019	<0.001	<0.001	0.003	<0.001	0.002	0.011	<0.001	0.001
	Validation	0.474	0.007	0.003	0.998	0.938	<0.001	0.049	<0.001	1.000	1.000	0.119	<0.001
Left tail test	Calibration	0.136	1.000	1.000	0.981	1.000	1.000	0.997	1.000	0.998	0.989	1.000	0.999
	Validation	0.526	0.993	0.997	0.002	0.062	1.000	0.951	1.000	<0.001	<0.001	0.881	1.000

watersheds. For these 6 watersheds, the JBC method performs better for most GCMs but not all. Over the 3 watersheds where both methods perform similarly (# 1, # 5 and # 11), relative performance depends on the GCM. For the remaining 3 watersheds, the IBC methods performs better for most GCMs. Reasons of these behaviors will be discussed later.

Performance is also dependent on the choice of hydrological metric. Figs. 9 and 10 show that the performance of the JBC method depends on hydrological metrics. Overall, the JBC method consistently performs better than the IBC method for most metrics for 10 out of 12 watersheds over the calibration period. For the Manic-5 watershed (# 1), the JBC method performs worse for mean and standard deviation of winter discharge, as well as some metrics about extreme flows. However, it performs better for other metrics. This is not surprising, as precipitation accumulates as snow over the winter, and, consequently, correlation between precipitation and temperature is less relevant to streamflow generation. For the Oueme watershed (# 12), the JBC method performs better for wet season metrics, but worse for dry season metrics indicating that complex methods may be detrimental in such cases. For the validation period, differences between both methods are small, with the exception of the Nation and Oueme watersheds. The JBC method is clearly better over 3 watersheds (# 3, # 6 and # 12), while the IBC method is better over watersheds # 4, # 9 and # 10.

5. Discussion and conclusion

Bias correction methods applied to climate model outputs usually act on each climate variable separately. Thus, there are concerns regarding whether inter-variable dependencies of climate model outputs are distorted and whether the bias of inter-variable dependencies affects the reliable assessment of climate change impacts. In this study, the impacts of correcting the P-T correlations of GCM simulations on hydrological modeling were investigated. Specifically, the performance of a JBC method was compared with the commonly used IBC method in terms of hydrological modeling over 12 watersheds located in various climate regimes.

The results show that GCMs are considerably biased in terms of modeling P-T dependencies. Even though the IBC method largely improves the simulation of individual variables simulated by GCMs, the inter-variable dependencies are mostly retained. Similar results were also obtained in other studies (e.g., Wilcke et al., 2013; Li et al., 2014). The JBC method performs reasonably well in reducing the biases of the GCM-simulated P-T correlation for the calibration period. However, for the validation period, its performance is monthly dependent. For example, the P-T correlation bias is considerably reduced in July, but not so much in November. The different performance of JBC between the calibration and validation periods is mostly due to the fact that the biases of the GCM P-T

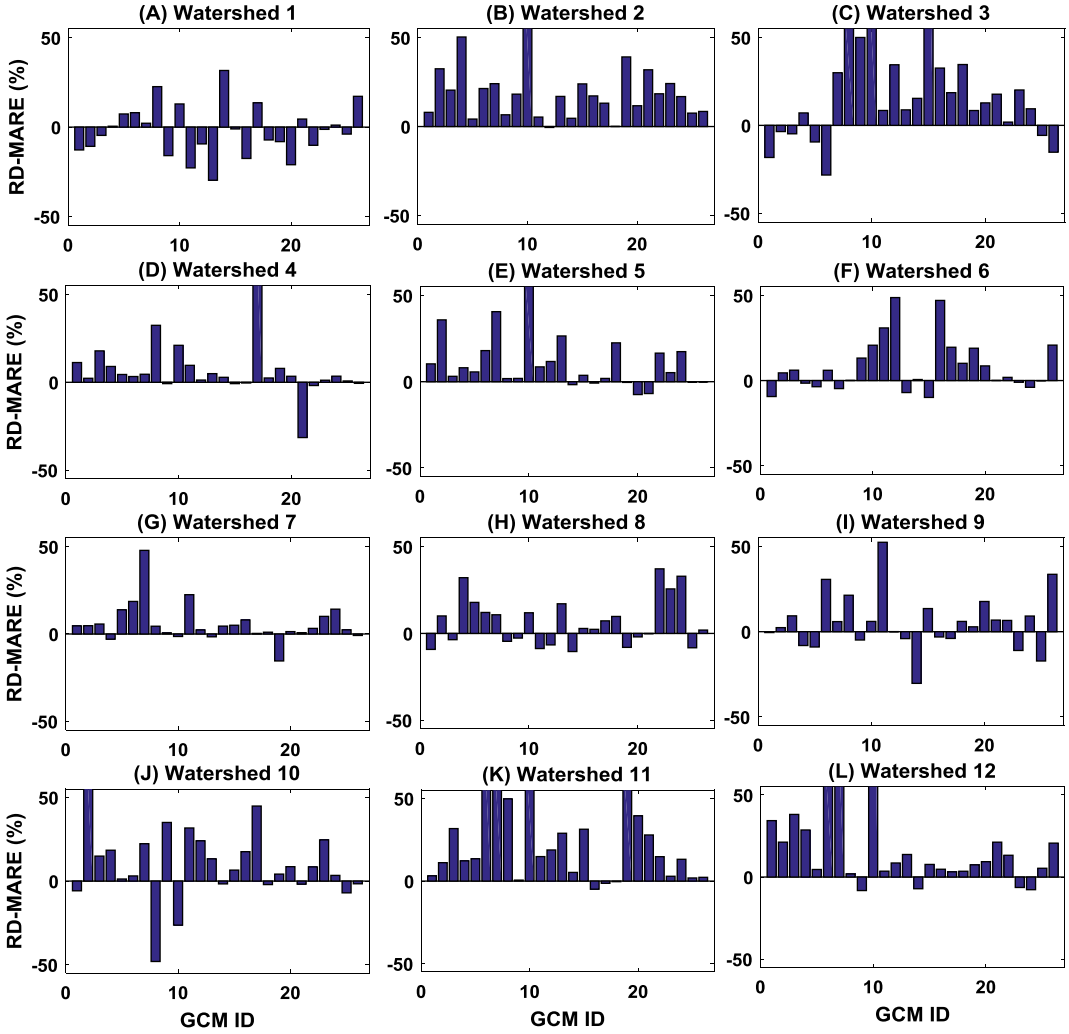

Fig. 7. The relative difference of mean absolute relative error (RD-MARE, %) between IBC- and JBC-corrected data simulated discharge for 26 GCMs and 12 watersheds over the calibration period.

correlation are not constant in time (more on this below). All bias correction methods are based on an assumption that climate model biases are stationary over time (Hewitson and Crane, 2006; Piani et al., 2010; Maraun, 2012; Maurer et al., 2013; Chen et al., 2015). However, this is not necessarily so in practice, even over two adjacent time periods, due to the natural climate variability (Chen et al., 2015). If the 'stationary assumption' is not verified, there is little hope that a bias correction method can work equally well over calibration and validation periods.

In terms of hydrological modeling, the JBC method performs significantly better than the IBC method for 11 out of the 12 watersheds over the calibration period. This highlights the importance of correcting the inter-variable dependencies of GCMs in hydrological modeling. For the coldest snowfall-dominated watershed (Manic-5 (# 1)), the advantage of JBC is not evident. This is as expected, because the water availability of this northern watershed is dominated by snowfall in the winter, as about 45% precipitation occurs as snow. Evaporation, which is determined by temperature, is not important for this watershed, as the temperature is relative low all the year round. The temperature mainly affects the snow melt time in the spring. In other words, the hydrology of the northern watershed may not be sensitive to the P-T dependencies, especially when taking into account the fact that the bias of the GCM P-T correlation is relatively small for this watershed.

Unlike for the calibration period, the JBC method does not perform consistently better than the IBC method for the validation period. In particular, it performs significantly better or comparably to the IBC method for 9 out of the 12 watersheds, but exhibit worse performance over the other 3 watersheds. The 6 watersheds where the JBC method exhibit advantage can be classified into two groups: arid or tropical watersheds (Carrot (# 2), Nation (# 3) and Oueme (# 12)) and snowfall-rainfall-mixed watersheds (Sacandaga (# 6) and Wolf (# 8)). For arid or tropical watersheds, the evaporation (determined by temperature) represents a considerable amount of water loss, and precipitation determines the available water that can be evaporated. Thus, the relationship between precipitation and temperature is tight. The tropical Oueme (# 12) watershed is a good example, over which the JBC method performs much better than the IBC method for most hydrological metrics. Even though the Umpqua watershed cannot be labeled as having an arid climate, its hydrological regimes nonetheless displays contrasting dry and wet seasons. The use of JBC is also important for snowfall-rainfall-mixed watersheds, as the temperature determines the proportion of precipitation that occurs as rain or snow, as well as the proportion of snow to melt (Li et al., 2014). However, for snowfall-dominated watersheds (Manic-5 (# 1), Yampa (# 9) and Yellowstone (# 10)), the advantage of the JBC method is not obvious. As mentioned earlier, this

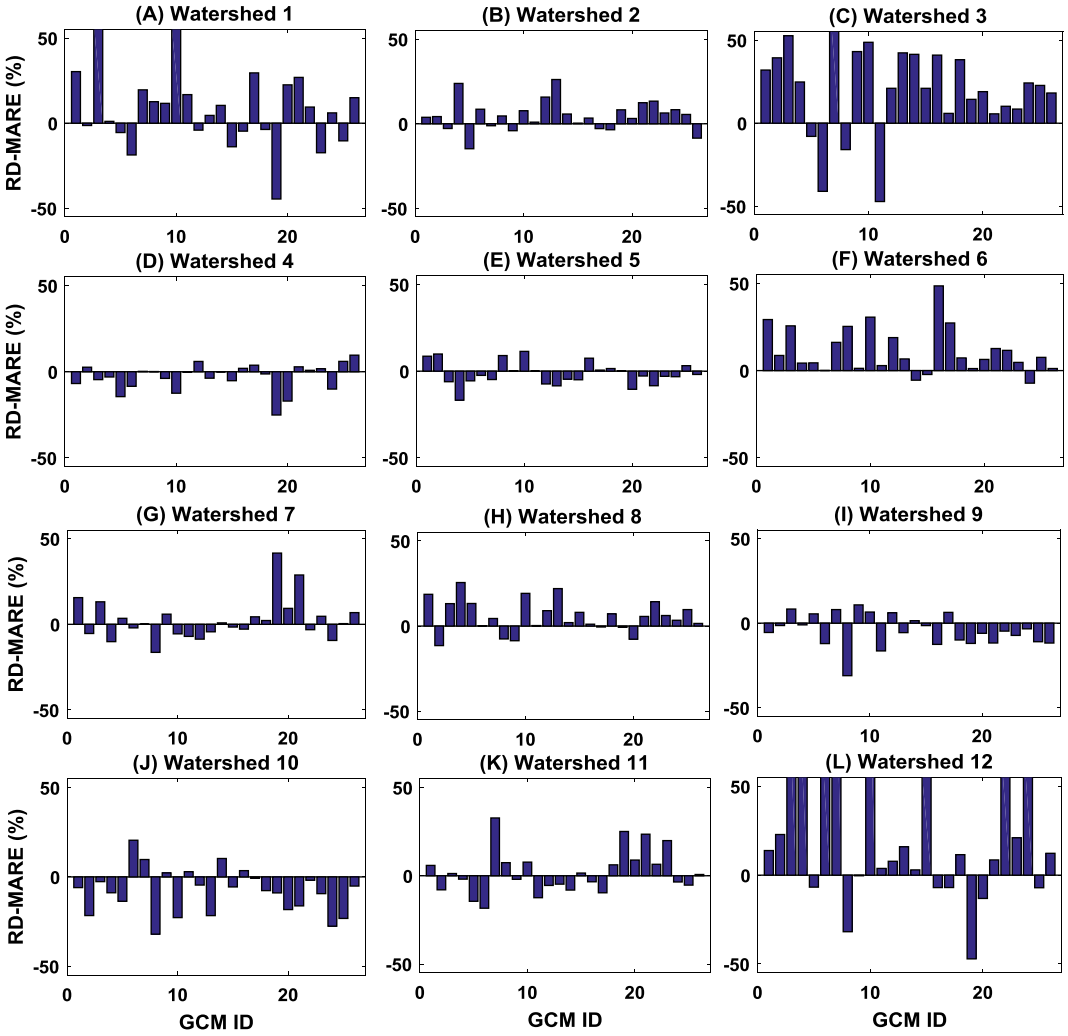

Fig. 8. The relative difference of mean absolute relative error (RD-MARE, %) between IBC- and JBC-corrected data simulated discharge for 26 GCMs and 12 watersheds over the validation period.

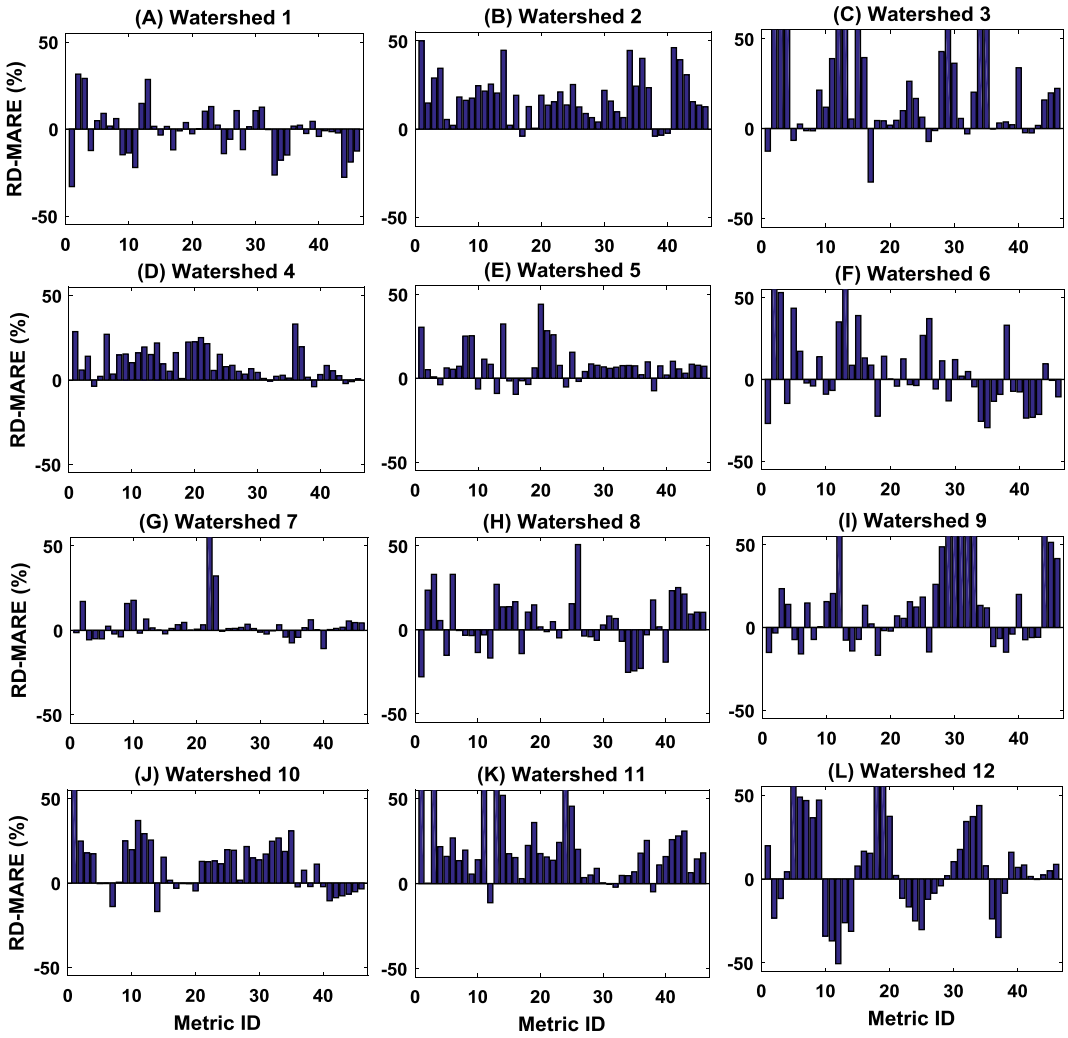

Fig. 9. The relative difference of mean absolute relative error (RD-MARE, %) between IBC- and JBC-corrected data simulated discharge for 46 hydrological metrics and 12 watersheds over the calibration period.

is because the temperature is very low for these watersheds in the winter, and a slight fluctuation in temperature would likely have little effect on snowmelt and evapotranspiration. Similarly, the advantage of the JBC method is also not significant for rainfall-dominated watersheds (Grand (# 5), Xiangjiang (# 11) and Chickasawhay (# 4)), due to the unsynchronized wet and heat seasons.

Besides the reasons mentioned above, the bias nonstationarity of GCM simulations may ultimately be the main culprit. Because of the nonstationary P-T correlations partly arisen from natural climate variability, the JBC method may have little advantage over the IBC method in terms of correcting the GCM P-T correlation, and may even induce additional biases for some months (e.g., November). This is especially true for the three watersheds where the IBC method performs better than its counterpart for hydrological simulations. To understand these results more clearly, the assumption of bias stationarity in P-T correlations is verified for the Yampa (# 9) and Yellowstone (# 10) watersheds. Fig. 11 presents the difference in biases between the validation and calibration periods (validation-calibration) for the monthly P-T correlations for all 26 GCMs. The observed monthly P-T correlations for all 12 watersheds are also presented for comparison (Fig. 12). The bias is defined as the difference in monthly P-T correlation between observations and GCM time series. The difference in bias is equivalent to the future bias remaining after a bias correction based on the calibration period has been performed. The results show that the bias of P-T correlations is clearly not stationary for these two watersheds, especially when taking into account the fact that the difference in bias between the calibration and validation periods is greater than the observed P-T correlations (lines 9 and 10 in Fig. 12). Due to the bias nonstationarity of the biases, the JBC method deteriorates the original P-T correlations simulated by GCMs. This explains why the IBC method performs significantly better than the JBC method for 3 out of 12 watersheds. Increasing the length of the calibration period so as to more comprehensively sample the natural climate variability may be able to improve the JBC performance to some extent.

In addition, even though all GCMs are capable of appropriately simulating the climate at the global scale, their performance vary at the regional and watershed scales, thus potentially explaining some of the differences.

The observed P-T correlation may be another factor which affects the performance of the JBC method. As presented in Fig. 12, correlations can be positive and negative, or even absent depending on the month. If the observed correlation is very small, a slight perturbation may result in large biases in P-T corrections,

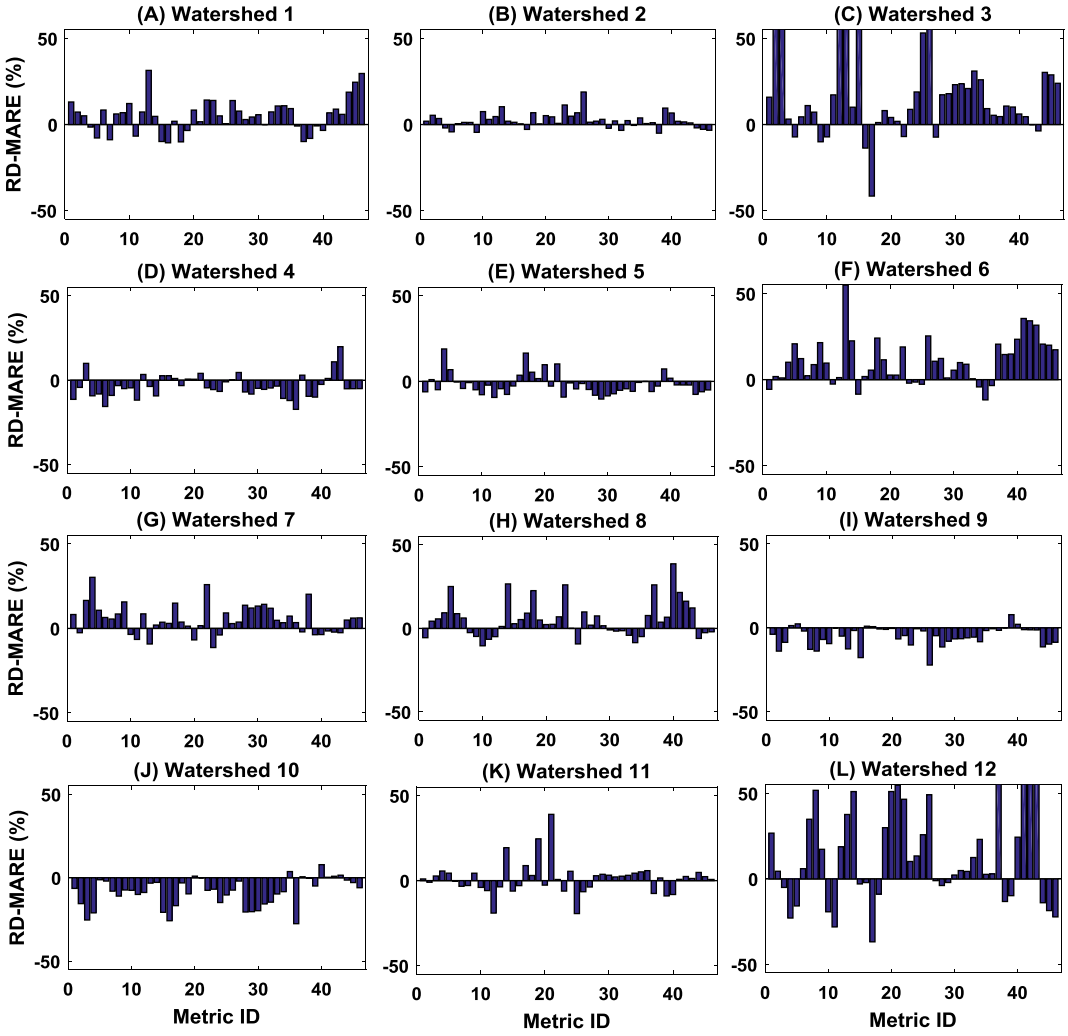

Fig. 10. The relative difference of mean absolute relative error (RD-MARE, %) between IBC- and JBC-corrected data simulated discharge for 46 hydrological metrics and 12 watersheds over the validation period.

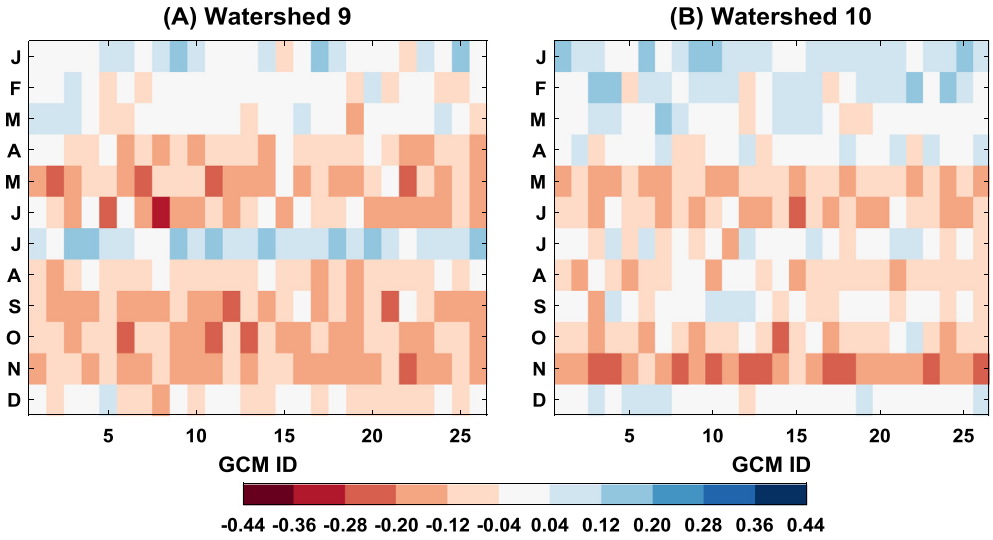

Fig. 11. Difference in biases of Kendall's tau correlation coefficients between calibration and validation periods for GCM-simulated precipitation and temperature over 12 months and 26 GCMs at the Yampa (A) and Yellowstone (B) watersheds.

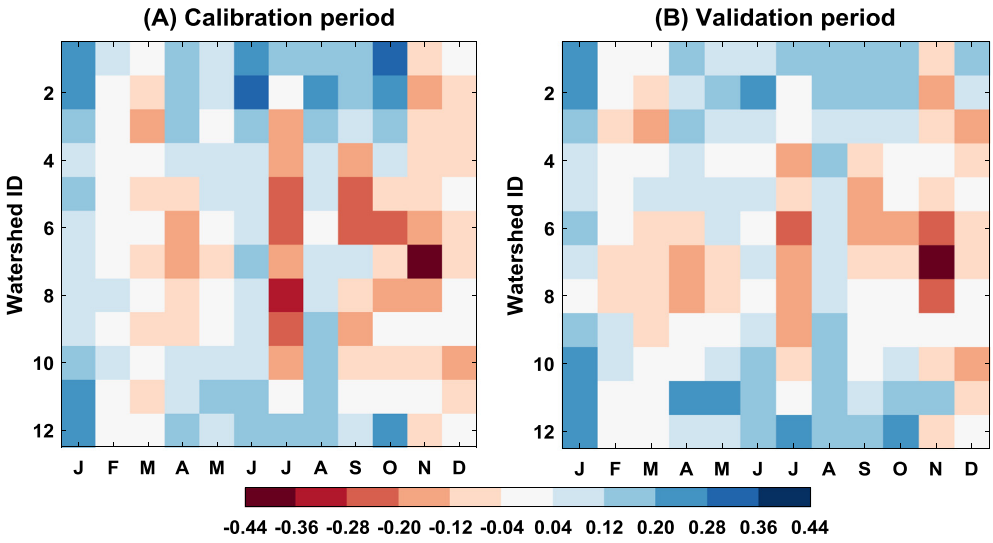

Fig. 12. Kendall's tau correlation coefficients between observed precipitation and temperature for 12 months and 12 watersheds.

especially when the bias of P-T corrections is not stationary. This may be why the JBC method can deteriorate the original P-T correlations in some cases.

Overall, this study emphasizes the importance of correcting the inter-variable dependencies of GCM simulations when using them for hydrological climate change impact studies. The JBC method is able to achieve this goal. Depending on hydrological simulations, the performance of the JBC method varies across watershed locations and hydrological metrics. The method may perform better than the IBC method for one hydrological metric, but may not for the other. This study only looked at the overall impacts of JBC using 46 hydrological metrics. For example, the paired *t*-test was conducted based on 46 pairs of hydrological metrics. A careful validation should always be performed for specific impact studies, for example, the impact of the JBC method on streamflow for specific seasons. In addition, the hydrological simulation carried out in the study was based on only one hydrological model, while the uncertainty of hydrological models was not considered. However, previous studies (e.g., Wilby and Harris, 2006; Chen et al., 2011a) have shown that the uncertainty due to hydrological models is much less important than that related to climate models. Therefore, it is unlikely that the conclusions drawn from this paper would change when using other hydrological models. Finally, this study only investigates the impact of JBC on hydrological modeling. The advantage of JBC may be more significant for other impact studies, such as those relating to agriculture, as the concurrent changes in precipitation and temperature are of greater interest than independent projections of each variable (Tebaldi and Lobell, 2008).

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (Grant No. 51779176, 51539009, 51339004), the Thousand Youth Talents Plan from the Organization Department of CCP Central Committee (Wuhan University, China), the Natural Science and Engineering Research Council of Canada (NSERC), Hydro-Québec, and the Ouranos Consortium on Regional Climatology and Adaption to Climate Change. Chao Li was also supported by the National Science Foundation under grant 1313897. The authors would like to acknowledge the contribution of the World Climate Research Program Working Group on Coupled Modelling, and all climate modeling groups listed in Table 2 for making available their respective model outputs.

References

Arsenault, R., Brissette, F., 2016. Multi-model averaging for continuous streamflow prediction in ungauged basins. Hydrol. Sci. J. 61 (13), 2443–2454.
Arsenault, R., Essou, G.R.C., Brissette, F.P., 2017. Improving hydrological model simulations with combined multi-input and multimodel averaging frameworks. J. Hydrol. Eng. 22 (4). https://doi.org/10.1061/(ASCE)HE.1943-5584.0001489.
Arsenault, R., Gatien, P., Renaud, B., Brissette, F., Martel, J.L., 2015. A comparative analysis of 9 multi-model averaging approaches in hydrological continuous streamflow simulation. J. Hydrol. 529, 754–767.
Arsenault, R., Poulin, A., Côté, P., Brissette, F., 2013. Comparison of stochastic optimization algorithms in hydrological model calibration. J. Hydrol. Eng. 19 (7), 1374–1384.
Buishand, T.A., Brandsma, T., 2001. Multisite simulation of daily precipitation and temperature in the Rhine Basin by nearest-neighbor resampling. Water Resour. Res. 37 (11), 2761–2776. https://doi.org/10.1002/2001WR000291.
Cannon, A.J., 2016. Multivariate bias correction of climate model output: matching marginal distributions and intervariable dependence structure. J. Clim. 29, 7045–7064.
Cannon, A.J., Sobie, S.R., Murdock, T.Q., 2015. Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? J. Clim. 28, 6938–6959.
Chen, J., Brissette, F.P., 2015. Combining stochastic weather Generation and ensemble weather forecasts for short-term streamflow prediction. Water Resour. Manag. 29 (9), 3329–3342.
Chen, J., Brissette, F.P., Lucas-Picher, P., Caya, D., 2017. Impacts of weighting climate models for hydro-meteorological climate change studies. J. Hydrol. 549, 534–546.
Chen, J., Brissette, F.P., Poulin, A., Leconte, R., 2011a. Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. Water Resour. Res. 47, W12509. https://doi.org/10.1029/2011WR010602.
Chen, J., Brissette, F.P., Leconte, R., 2011b. Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. J. Hydrol. 401, 190–202.
Chen, J., Brissette, F.P., Chaumont, D., Braun, M., 2013a. Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. Water Resour. Res. 49, 4187–4205. https://doi.org/10.1002/wrcr.20331.
Chen, J., Brissette, F.P., Chaumont, D., Braun, M., 2013b. Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate change impacts on hydrology over two North American river basins. J. Hydrol. 479 (4), 200–214.
Chen, J., Brissette, F.P., Lucas-Picher, P., 2015. Assessing the limits of bias-correcting climate model outputs for climate change impact studies. J. Geophys. Res. Atmos. 120. https://doi.org/10.1002/2014JD022635.
Chen, J., St-Denis, B.G., Brissette, F.P., Lucas-Picher, P., 2016. Using natural variability as a baseline to evaluate the performance of bias correction methods in hydrological climate change impact studies. J. Hydrometeorol. 17 (8), 2155–2174.
Duan, Q., Schaake, J., Andreassian, V., Franks, S., Goteti, G., Gupta, H.V., Gusev, Y.M., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O.N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., Wood, E.F., 2006. Model Parameter Estimation Experiment (MOPEX): an overview of

science strategy and major results from the second and third workshops. J. Hydrol. 320, 3–17.
Essou, G.R.C., Brissette, F., 2013. Climate Change Impacts on the Ouémé River, Benin, West Africa. J, Earth Sci. Clim. Change 4, 161. https://doi.org/10.4172/2157-7617.1000161.
Gennaretti, F., Sangelantoni, L., Grenier, F., 2015. Toward daily climate scenarios for Canadian Arctic coastal zones with more realistic temperature-precipitation interdependence. J. Geophys. Res. Atms. 120, 11862–11877. https://doi.org/10.1002/2015JD023890.
Gleckler, P.J., Taylor, K.E., Doutriaux, C., 2008. Performance metrics for climate models. J. Geophys. Res. 113, D06104. https://doi.org/10.1029/2007JD008972.
Hansen, N., Ostermeier, A., 1996. Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. paper presented at 1996 IEEE International Conference on Evolutionary Computation, Nagoya, Japan, IEEE Neural Network council, and Society of Instrument and Control Engineers.
Hansen, N., Ostermeier, A., 2001. Completely derandomized self-adaptation in evolution strategies. Evol. Comput. 9 (2), 159–195.
Hewitson, B.C., Crane, R.G., 2006. Consensus between GCM climate change projections with empirical downscaling: Precipitation downscaling over South Africa. Int. J. Climatol. 26, 1315–1337.
Hutchinson, M.F., McKenney, D.W., Lawrence, K., Pedlar, J.H., Hopkinson, R.F., Milewska, E., Papadopol, P., 2009. Development and testing of Canada-wide interpolated spatial models of daily minimum–maximum temperature and precipitation for 1961–2003. B. Am. Meteorol. Soc. 48, 725–741.
Immerzeel, W.W., Peterson, L., Ragettli, S., Pellicciotti, F., 2014. The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas. Water Resour. Res. 50, 2212–2226. https://doi.org/10.1002/2013WR014506.
Ines, A.V.M., Hansen, J.W., 2006. Bias correction of daily GCM rainfall for crop simulation studies. Agric. For. Meteorol. 138 (1–4), 44–53. https://doi.org/10.1016/j.agrformet.2006.03.009.
Kendall, M.G., 1970. Rank Correlation Methods. London, Griffin.
Li, C., Sinha, E., Horton, D.E., Diffenbaugh, N.S., Michalak, A.M., 2014. Joint bias correction of temperature and precipitation in climate model simulations. J. Geophys. Res. Atmos. 119 (23), 13153–13162.
Livneh, B., Rosenberg, E.A., Lin, C., Mishra, V., Andreadis, K.M., Maurer, E.P., Lettenmaier, D.P., 2013. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous united states: update and extensions. J. Climate 26, 9384–9392.
Maraun, D., 2012. Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. Geophys. Res. Lett. 39, L06706. https://doi.org/10.1029/2012GL051210.
Martel, J.-L., Demeester, K., Brissette, F., Poulin, A., Arsenault, R., 2017. HMETS – a simple and efficient hydrology model for teaching hydrological modelling, flow forecasting and climate change impacts. Int. J. Eng. Educ. 33 (4), 1307–1316.
Maurer, E.P., Wood, A.W., Adam, J.C., Lettenmaier, D.P., Nijssen, B., 2002. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous united states. J. Clim. 15, 3237–3251.
Maurer, E.P., Hidalgo, H.G., Das, T., Dettinger, M.D., Cayan, D.R., 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol. Earth Syst. Sci. 14, 1125–1138.
Maurer, E.P., Das, T., Cayan, D.R., 2013. Errors in climate model daily precipitation and temperature output: Time invariance and implications for bias correction. Hydrol. Earth Syst. Sci. 17, 2147–2159.
Minville, M., Brissette, F., Leconte, R., 2008. Uncertainty of the impact of climate change on the hydrology of a nordic watershed. J. Hydrol. 358, 70–83.
Mueller, B., Seneviratne, S.I., 2014. Systematic land climate and evapotranspiration biases in CMIP5 simulations. Geophys. Res. Lett. 40, 128–134. https://doi.org/10.1002/2013GL058055.
Nash, J.E., Sutcliffe, W.H., 1970. River flow forecasting through conceptual models: Part 1. A discussion of principles. J. Hydrol. 10 (3), 282–290.
Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andreassian, V., Anctil, F., Loumagne, C., 2005. Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2–Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling. J. Hydrol. 303, 290–306.
Piani, C., Haerter, J.O., 2012. Two dimensional bias correction of temperature and precipitation copulas in climate models. Geophys. Res. Lett. 39, L20401. https://doi.org/10.1029/2012GL053839.
Piani, C., Haerter, O., Corpola, E., 2010. Statistical bias correction for daily precipitation in regional climate models over Europe. Theor. Appl. Climatol. 99, 187–192.
Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., Participants, D.M.I.P., 2004. Overall distributed model intercomparison project results. J. Hydrol. 298 (1–4), 27–60.
Rhynsburger, D., 1973. Analytic delineation of thiessen polygons. Geographical Anal. 5, 133–144.
Schmidli, J., Frei, C., Vidale, P.L., 2006. Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. Int. J. Climatol. 26, 679–689.
Sharma, D., Das Gupta, A., Babel, M.S., 2007. Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping river basin. Thailand. Hydrol. Earth Syst. Sci. 11 (4), 1373–1390.
Smith, M.B., Koren, V., Zhang, Z., Zhang, Y., Reed, S.M., Cui, Z., Anderson, E., 2012. Results of the DMIP 2 Oklahoma experiments. J Hydrol. 418–419, 17–48.
Tebaldi, C., Lobell, D.B., 2008. Toward probabilistic projections of climate change impacts on global crop yields. Geophys. Res. Lett. 35, L08705. https://doi.org/10.1029/2008GL033423.
Teutschbein, C., Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J. Hydrol. 456–457, 12–29. https://doi.org/10.1016/j.jhydrol.2012.05.052.
Teutschbein, C., Seibert, J., 2013. Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? Hydrol. Earth Syst. Sci. 17, 5061–5077.
Thrasher, B., Maurer, E.P., McKellar, C., Duffy, P.V., 2012. Technical note: bias correcting climate model simulated daily temperature extremes with quantile mapping. Hydrol. Earth Syst. Sci. 16, 3309–3314.
USGS - Subcommittee, H., 1982. Guidelines for determining flood flow frequency. Bulletin B, 17.
Vehvilainen, B., 1992. Snow Cover Models in Operational Watershed Forecasting PhD diss. National Board of Waters and the Environment, Helsinki.
Vrac, M., Friederichs, P., 2015. Multivariate-intervariable, spatial, and temporal-bias correction. J. Clim. 28, 218–237.
Wilby, R.L., Harris, I., 2006. A framework for assessing uncertainties in climate change impacts: low-flow scenarios for the River Thames, UK. Water Resour. Res. 42, W02419. https://doi.org/10.1029/2005WR004065.
Wilcke, R.A.I., Mendlik, T., Gobiet, A., 2013. Multi-variable error correction of regional climate models. Climatic Change 120 (4), 871–887.